

The modern tools of quantum mechanics

A tutorial on quantum states, measurements, and operations

Matteo G A Paris^{1,2,3}

¹ Dipartimento di Fisica dell'Università degli Studi di Milano, I-20133 Milano, Italia, EU

² CNISM - Udr Milano, I-20133 Milano, Italia, EU.

³ e-mail: matteo.paris@fisica.unimi.it

Abstract. We address the basic postulates of quantum mechanics and point out that they are formulated for a closed isolated system. Since we are mostly dealing with systems that interact or have interacted with the rest of the universe one may wonder whether a suitable modification is needed, or in order. This is indeed the case and this tutorial is devoted to review the modern tools of quantum mechanics, which are suitable to describe states, measurements, and operations of realistic, not isolated, systems. We underline the central role of the Born rule and and illustrate how the notion of density operator naturally emerges, together the concept of purification of a mixed state. In reexamining the postulates of standard quantum measurement theory, we investigate how they may formally generalized, going beyond the description in terms of selfadjoint operators and projective measurements, and how this leads to the introduction of generalized measurements, probability operator-valued measures (POVMs) and detection operators. We then state and prove the Naimark theorem, which elucidates the connections between generalized and standard measurements and illustrates how a generalized measurement may be physically implemented. The "impossibility" of a joint measurement of two non commuting observables is revisited and its canonical implementation as a generalized measurement is described in some details. The notion of generalized measurement is also used to point out the heuristic nature of the so-called Heisenberg principle. Finally, we address the basic properties, usually captured by the request of unitarity, that a map transforming quantum states into quantum states should satisfy to be physically admissible, and introduce the notion of complete positivity (CP). We then state and prove the Stinespring/Kraus-Choi-Sudarshan dilation theorem and elucidate the connections between the CP-maps description of quantum operations, together with their operator-sum representation, and the customary unitary description of quantum evolution. We also address transposition as an example of positive map which is not completely positive, and provide some examples of generalized measurements and quantum operations.

Contents

1	Introduction	2
2	Quantum states	4
2.1	Density operator and partial trace	4
2.1.1	Conditional states	6
2.2	Purity and purification of a mixed state	6
3	Quantum measurements	7
3.1	Probability operator-valued measure and detection operators	8
3.2	The Naimark theorem	9
3.2.1	Conditional states in generalized measurements	13
3.3	Joint measurement of non commuting observables	13
3.4	About the so-called Heisenberg principle	15
3.5	The quantum roulette	16
4	Quantum operations	17
4.1	The operator-sum representation	18
4.1.1	The dual map and the unitary equivalence	20
4.2	The random unitary map and the depolarizing channel	21
4.3	Transposition and partial transposition	22
5	Conclusions	23
A	Trace and partial trace	24
B	Uncertainty relations	25

1 Introduction

Quantum information science is a novel discipline which addresses how quantum systems may be exploited to improve the processing, transmission, and storage of information. This field has fostered new experiments and novel views on the conceptual foundations of quantum mechanics, and also inspired much current research on coherent quantum phenomena, with quantum optical systems playing a prominent role. Yet, the development of quantum information had so far little impact on the way that quantum mechanics is taught, both at graduate and undergraduate levels. This tutorial is devoted to review the mathematical tools of quantum mechanics and to present a modern reformulation of the basic postulates which is suitable to describe quantum systems in interaction with their environment, and with any kind of measuring and processing devices.

We use Dirac notation throughout this tutorial and by *system* we mean a single given degree of freedom (spin, position, angular momentum,...) of a physical entity. The postulates of quantum mechanics are a list of prescriptions to summarize

1. how we describe the states of a physical system;
2. how we describe the measurements performed on a physical system;
3. how we describe the evolution of a physical system, either because of the dynamics or due to a measurement.

In this section we present a picoreview of the basic postulates of quantum mechanics in order to introduce notation and point out both i) the implicit assumptions contained in the standard formulation, and ii) the need of a reformulation in terms of more general mathematical objects. For our purposes the postulates of quantum mechanics may be grouped and summarized as follows

Postulate 1 (States of a quantum system) The possible states of a physical system correspond to normalized vectors $|\psi\rangle$, $\langle\psi|\psi\rangle = 1$, of a Hilbert space H . Composite systems, either made by more than one physical object or by the different degrees of

freedom of the same entity, are described by tensor product $H_1 \otimes H_2 \otimes \dots$ of the corresponding Hilbert spaces, and the overall state of the system is a vector in the global space. As far as the Hilbert space description of physical systems is adopted, then we have the *superposition principle*, which says that if $|\psi_1\rangle$ and $|\psi_2\rangle$ are possible states of a system, then also any (normalized) linear combination $\alpha|\psi_1\rangle + \beta|\psi_2\rangle$, $\alpha, \beta \in \mathbb{C}$, $|\alpha|^2 + |\beta|^2 = 1$ of the two states is an admissible state of the system.

Postulate 2 (Quantum measurements) Observable quantities are described by Hermitian operators X . Any hermitian operator $X = X^\dagger$, admits a spectral decomposition $X = \sum_x x P_x$, in terms of its real eigenvalues x , which are the possible value of the observable, and of the projectors $P_x = |x\rangle\langle x|$, $P_x, P_{x'} = \delta_{xx'} P_x$ on its eigenvectors $X|x\rangle = x|x\rangle$, which form a basis for the Hilbert space, i.e. a complete set of orthonormal states with the properties $\langle x|x'\rangle = \delta_{xx'}$ (orthonormality), and $\sum_x |x\rangle\langle x| = \mathbb{I}$ (completeness, we omitted to indicate the dimension of the Hilbert space). The probability of obtaining the outcome x from the measurement of the observable X is given by $p_x = |\langle\psi|x\rangle|^2$, i.e

$$p_x = \langle\psi|P_x|\psi\rangle = \sum_n \langle\psi|\varphi_n\rangle\langle\varphi_n|P_x|\psi\rangle = \sum_n \langle\varphi_n|P_x|\psi\rangle\langle\psi|\varphi_n\rangle = \text{Tr} [|\psi\rangle\langle\psi| P_x] , \quad (1)$$

and the overall expectation value by

$$\langle X \rangle = \langle\psi|X|\psi\rangle = \text{Tr} [|\psi\rangle\langle\psi| X] .$$

This is the *Born rule*, which represents the fundamental recipe to connect the mathematical description of a quantum state to the prediction of quantum theory about the results of an actual experiment. The state of the system *after* the measurement is the (normalized) projection of the state *before* the measurement on the eigenspace of the observed eigenvalue, i.e.

$$|\psi_x\rangle = \frac{1}{\sqrt{p_x}} P_x |\psi\rangle .$$

Postulate 3 (Dynamics of a quantum system) The dynamical evolution of a physical system is described by unitary operators: if $|\psi_0\rangle$ is the state of the system at time t_0 then the state of the system at time t is given by $|\psi_t\rangle = U(t, t_0)|\psi_0\rangle$, with $U(t, t_0)U^\dagger(t, t_0) = U^\dagger(t, t_0)U(t, t_0) = \mathbb{I}$. For systems governed by a time-independent Hamiltonian $H \neq H(t)$ we have $U(t, t_0) = \exp\{-iH(t - t_0)\}$.

We will denote by $L(H)$ the linear space of (linear) operators from H to H , which itself is a Hilbert space with scalar product provided by the trace operation, i.e. upon denoting by $|A\rangle\rangle$ operators seen as elements of $L(H)$, we have $\langle\langle A|B\rangle\rangle = \text{Tr}[A^\dagger B]$ (see Appendix A for details on the trace operation). In the following, we are going to address both finite- and infinite-dimensional systems, but despite this fact we are not going to pay attention to the domain of operators, and to all the subtleties related to the geometry of infinite-dimensional Hilbert spaces.

As it is apparent from their formulation, the postulates of quantum mechanics, as reported above, are about a closed isolated system. On the other hand, we are mostly dealing with system that interacts or have interacted with the rest of the universe, either during their dynamical evolution, or when subjected to a measurement. As a consequence, one may wonder whether a suitable modification is needed, or in order. This is indeed the case and the rest of his tutorial is devoted to review the mathematical tools of quantum mechanics and to present a modern reformulation of the basic postulates which is suitable to describe, design and control quantum systems in interaction with their environment, and with any kind of measuring and processing devices.

2 Quantum states

2.1 Density operator and partial trace

Suppose to have a quantum system whose preparation is not completely under control. What we know is that the system is prepared in the state $|\psi_k\rangle$ with probability p_k , i.e. that the system is described by the statistical ensemble $\{p_k, |\psi_k\rangle\}$, $\sum_k p_k = 1$, where the states $\{|\psi_k\rangle\}$ are not, in general, orthogonal. The expected value of an observable X may be evaluated as follows

$$\begin{aligned}\langle X \rangle &= \sum_k p_k \langle X \rangle_k = \sum_k p_k \langle \psi_k | X | \psi_k \rangle = \sum_{n,p,k} p_k \langle \psi_k | \varphi_n \rangle \langle \varphi_n | X | \varphi_p \rangle \langle \varphi_p | \psi_k \rangle \\ &= \sum_{n,p,k} p_k \langle \varphi_p | \psi_k \rangle \langle \psi_k | \varphi_n \rangle \langle \varphi_n | X | \varphi_p \rangle = \sum_{n,p} \langle \varphi_p | \varrho | \varphi_n \rangle \langle \varphi_n | X | \varphi_p \rangle \\ &= \sum_p \langle \varphi_p | \varrho X | \varphi_p \rangle = \text{Tr} [\varrho X]\end{aligned}$$

where

$$\varrho = \sum_k p_k |\psi_k\rangle \langle \psi_k|$$

is the *statistical (density) operator* describing the system under investigation. The $|\varphi_n\rangle$'s in the above formula are a basis for the Hilbert space, and we used the trick of suitably inserting two resolutions of the identity $\mathbb{I} = \sum_n |\varphi_n\rangle \langle \varphi_n|$. The formula is of course trivial if the $|\psi_k\rangle$'s are themselves a basis or a subset of a basis.

Theorem 1 (Density operator) *An operator ϱ is the density operator associated to an ensemble $\{p_k, |\psi_k\rangle\}$ if and only if it is a positive $\varrho > 0$ (hence selfadjoint) operator with unit trace $\text{Tr} [\varrho] = 1$.*

Proof : If $\varrho = \sum_k p_k |\psi_k\rangle \langle \psi_k|$ is a density operator then $\text{Tr}[\varrho] = \sum_k p_k = 1$ and for any vector $|\varphi\rangle \in H$, $\langle \varphi | \varrho | \varphi \rangle = \sum_k p_k |\langle \varphi | \psi_k \rangle|^2 \geq 0$. Viceversa, if ϱ is a positive operator with unit trace then it can be diagonalized and the sum of eigenvalues is equal to one. Thus it can be naturally associated to an ensemble. \square

As it is true for any operator, the density operator may be expressed in terms of its matrix elements in a given basis, i.e. $\varrho = \sum_{n,p} \varrho_{np} |\varphi_n\rangle \langle \varphi_p|$ where $\varrho_{np} = \langle \varphi_n | \varrho | \varphi_p \rangle$ is usually referred to as the *density matrix* of the system. Of course, the density matrix of a state is diagonal if we use a basis which coincides or includes the set of eigenvectors of the density operator, otherwise it contains off-diagonal elements. Different ensembles may lead to the same density operator. In this case they have the same expectation values for any operator and thus are physically indistinguishable. In other words, different ensembles leading to the same density operator are actually the same state, i.e. the density operator provides the natural and most fundamental quantum description of physical systems.

How this reconciles with Postulate 1 dictating that *physical systems are described by vectors in a Hilbert space*?

In order to see how it works let us first notice that, according to the postulates reported above, the action of "measuring nothing" should be described by the identity operator \mathbb{I} . Indeed the identity is Hermitian and has the single eigenvalues 1, corresponding to the persistent result of measuring nothing. Besides, the eigenprojector corresponding to the eigenvalue 1 is the projector over the whole Hilbert space and thus we have the consistent prediction that the state after the "measurement" is

left unchanged. Let us now consider a situation in which a bipartite system prepared in the state $|\psi_{AB}\rangle \in H_A \otimes H_B$ is subjected to the measurement of an observable $X = \sum_x P_x \in L(H_A)$, $P_x = |x\rangle\langle x|$ i.e. a measurement involving only the degree of freedom described by the Hilbert space H_A . The overall observable measured on the global system is thus $\mathbf{X} = X \otimes \mathbb{I}_B$, with spectral decomposition $\mathbf{X} = \sum_x x \mathbf{Q}_x$, $\mathbf{Q}_x = P_x \otimes \mathbb{I}_B$. The probability distribution of the outcomes is then obtained using the Born rule, i.e.

$$p_x = \text{Tr}_{AB} [|\psi_{AB}\rangle\langle\psi_{AB}| P_x \otimes \mathbb{I}_B]. \quad (2)$$

On the other hand, since the measurement has been performed on the sole system A , one expects the Born rule to be valid also at the level of the single system A , and a question arises on the form of the object ϱ_A which allows one to write $p_x = \text{Tr}_A [\varrho_A P_x]$ i.e. the Born rule as a trace only over the Hilbert space H_A . Upon inspecting Eq. (2) one sees that a suitable mapping $|\psi_{AB}\rangle\langle\psi_{AB}| \rightarrow \varrho_A$ is provided by the partial trace $\varrho_A = \text{Tr}_B [|\psi_{AB}\rangle\langle\psi_{AB}|]$. Indeed, for the operator ϱ_A defined as the partial trace, we have $\text{Tr}_A [\varrho_A] = \text{Tr}_{AB} [|\psi_{AB}\rangle\langle\psi_{AB}|] = 1$ and, for any vector $|\varphi\rangle \in H_A$, $\langle\varphi_A|\varrho_A|\varphi_A\rangle = \text{Tr}_{AB} [|\psi_{AB}\rangle\langle\psi_{AB}| |\varphi_A\rangle\langle\varphi_A| \otimes \mathbb{I}_B] \geq 0$. Being a positive, unit trace, operator ϱ_A is itself a density operator according to Theorem 1. As a matter of fact, the partial trace is the unique operation which allows to maintain the Born rule at both levels, i.e. the unique operation leading to the correct description of observable quantities for subsystems of a composite system. Let us state this as a little theorem [1]

Theorem 2 (Partial trace) *The unique mapping $|\psi_{AB}\rangle\langle\psi_{AB}| \rightarrow \varrho_A = f(\psi_{AB})$ from $H_A \otimes H_B$ to H_A for which $\text{Tr}_{AB} [|\psi_{AB}\rangle\langle\psi_{AB}| P_x \otimes \mathbb{I}_B] = \text{Tr}_A [f(\psi_{AB}) P_x]$ is the partial trace $f(\psi_{AB}) \equiv \varrho_A = \text{Tr}_B [|\psi_{AB}\rangle\langle\psi_{AB}|]$.*

Proof Basically the proof reduces to the fact that the set of operators on H_A is itself a Hilbert space $L(H_A)$ with scalar product given by $\langle\langle A|B\rangle\rangle = \text{Tr}[A^\dagger B]$. If we consider a basis of operators $\{M_k\}$ for $L(H_A)$ and expand $f(\psi_{AB}) = \sum_k M_k \text{Tr}_A [M_k^\dagger f(\psi_{AB})]$, then since the map f has to preserve the Born rule, we have

$$f(\psi_{AB}) = \sum_k M_k \text{Tr}_A [M_k^\dagger f(\psi_{AB})] = \sum_k M_k \text{Tr}_{AB} [M_k^\dagger \otimes \mathbb{I}_B |\psi_{AB}\rangle\langle\psi_{AB}|]$$

and the thesis follows from the fact that in a Hilbert space the decomposition on a basis is unique. \square

The above result can be easily generalized to the case of a system which is initially described by a density operator ϱ_{AB} , and thus we conclude that when we focus attention to a subsystem of a composite larger system the unique mathematical description of the act of ignoring part of the degrees of freedom is provided by the partial trace. It remains to be proved that the partial trace of a density operator is a density operator too. This is a very consequence of the definition that we put in the form of another little theorem.

Theorem 3 *The partial traces $\varrho_A = \text{Tr}_B [\varrho_{AB}]$, $\varrho_B = \text{Tr}_A [\varrho_{AB}]$ of a density operator ϱ_{AB} of a bipartite system, are themselves density operators for the reduced systems.*

Proof We have $\text{Tr}_A [\varrho_A] = \text{Tr}_B [\varrho_B] = \text{Tr}_{AB} [\varrho_{AB}] = 1$ and, for any state $|\varphi_A\rangle \in H_A$, $|\varphi_B\rangle \in H_B$,

$$\begin{aligned} \langle\varphi_A|\varrho_A|\varphi_A\rangle &= \text{Tr}_{AB} [\varrho_{AB} |\varphi_A\rangle\langle\varphi_A| \otimes \mathbb{I}_B] \geq 0 \\ \langle\varphi_B|\varrho_B|\varphi_B\rangle &= \text{Tr}_{AB} [\varrho_{AB} \mathbb{I}_A \otimes |\varphi_B\rangle\langle\varphi_B|] \geq 0. \quad \square \end{aligned}$$

2.1.1 Conditional states

From the above results it also follows that when we perform a measurement on one of the two subsystems, the state of the "unmeasured" subsystem after the observation of a specific outcome may be obtained as the partial trace of the overall post measurement state, i.e. the projection of the state before the measurement on the eigenspace of the observed eigenvalue, in formula

$$\varrho_{Bx} = \frac{1}{p_x} \text{Tr}_A [P_x \otimes \mathbb{I}_B \varrho_{AB} P_x \otimes \mathbb{I}_B] = \frac{1}{p_x} \text{Tr}_A [\varrho_{AB} P_x \otimes \mathbb{I}_B] \quad (3)$$

where, in order to write the second equality, we made use of the circularity of the trace (see Appendix A) and of the fact that we are dealing with a factorized projector. The state ϱ_{Bx} will be also referred to as the "conditional state" of system B after the observation of the outcome x from a measurement of the observable X performed on the system A .

Exercise 1 Consider a bidimensional system (say the spin state of a spin $\frac{1}{2}$ particle) and find two ensembles corresponding to the same density operator.

Exercise 2 Consider a spin $\frac{1}{2}$ system and the ensemble $\{p_k, |\psi_k\rangle\}$, $k = 0, 1$, $p_0 = p_1 = \frac{1}{2}$, $|\psi_0\rangle = |0\rangle$, $|\psi_1\rangle = |1\rangle$, where $|k\rangle$ are the eigenstates of σ_3 . Write the density matrix in the basis made of the eigenstates of σ_3 and then in the basis of σ_1 . Then, do the same but for the ensemble obtained from the previous one by changing the probabilities to $p_0 = \frac{1}{4}$, $p_1 = \frac{3}{4}$.

Exercise 3 Write down the partial traces of the state $|\psi\rangle\rangle = \cos\phi|00\rangle\rangle + \sin\phi|11\rangle\rangle$, where we used the notation $|jk\rangle\rangle = |j\rangle \otimes |k\rangle$.

2.2 Purity and purification of a mixed state

As we have seen in the previous section when we observe a portion, say A , of a composite system described by the vector $|\psi_{AB}\rangle\rangle \in H_A \otimes H_B$, the mathematical object to be inserted in the Born rule in order to have the correct description of observable quantities is the partial trace, which individuates a density operator on H_A . Actually, also the converse is true, i.e. any density operator on a given Hilbert space may be viewed as the partial trace of a state vector on a larger Hilbert space. Let us prove this constructively: if ϱ is a density operator on H , then it can be diagonalized by its eigenvectors and it can be written as $\varrho = \sum_k \lambda_k |\psi_k\rangle\langle\psi_k|$; then we introduce another Hilbert space K , with dimension at least equal to the number of nonzero eigenvalues of ϱ and a basis $\{|\theta_k\rangle\}$ in K , and consider the vector $|\varphi\rangle\rangle \in H \otimes K$ given by $|\varphi\rangle\rangle = \sum_k \sqrt{\lambda_k} |\psi_k\rangle \otimes |\theta_k\rangle$. Upon tracing over the Hilbert space K , we have

$$\text{Tr}_K [|\varphi\rangle\rangle\langle\langle\varphi|] = \sum_{kk'} \sqrt{\lambda_k \lambda_{k'}} |\psi_k\rangle\langle\psi_{k'}| \langle\theta_{k'}|\theta_k\rangle = \sum_k \lambda_k |\psi_k\rangle\langle\psi_k| = \varrho.$$

Any vector on a larger Hilbert space which satisfies the above condition is referred to as a *purification* of the given density operator. Notice that, as it is apparent from the proof, there exist infinite purifications of a density operator. Overall, putting together this fact with the conclusions from the previous section, we are led to reformulate the first postulate to say that *quantum states of a physical system are described by density operators*, i.e. positive operators with unit trace on the Hilbert space of the system.

A suitable measure to quantify how far a density operator is from a projector is the so-called *purity*, which is defined as the trace of the square density operator $\mu[\varrho] = \text{Tr}[\varrho^2] = \sum_k \lambda_k^2$, where the λ_k 's are the eigenvalues of ϱ . Density operators made by a projector $\varrho = |\psi\rangle\langle\psi|$ have $\mu = 1$ and are referred to as *pure states*, whereas for any $\mu < 1$ we have a *mixed state*. Purity of a state ranges in the interval $1/d \leq \mu \leq 1$ where d is the dimension of the Hilbert space. The lower bound is found looking for the minimum of $\mu = \sum_k \lambda_k^2$ with the constraint $\sum_k \lambda_k = 1$, and amounts to minimize the function $F = \mu + \gamma \sum_k \lambda_k$, γ being a Lagrange multipliers. The solution is $\lambda_k = 1/d$, $\forall k$, i.e. the maximally mixed state $\varrho = \mathbb{I}/d$, and the corresponding purity is $\mu = 1/d$.

When a system is prepared in a pure state we have the maximum possible information on the system according to quantum mechanics. On the other hand, for mixed states the degree of purity is connected with the amount of information we are missing by looking at the system only, while ignoring the *environment*, i.e. the rest of the universe. In fact, by looking at a portion of a composite system we are ignoring the information encoded in the correlations between the portion under investigation and the rest of system: This results in a smaller amount of information about the state of the subsystem itself. In order to emphasize this aspect, i.e. the existence of residual ignorance about the system, the degree of mixedness may be quantified also by the Von Neumann (VN) entropy $S[\varrho] = -\text{Tr}[\varrho \log \varrho] = -\sum_n \lambda_n \log \lambda_n$, where $\{\lambda_n\}$ are the eigenvalues of ϱ . We have $0 \leq S[\varrho] \leq \log d$: for a pure state $S[|\psi\rangle\langle\psi|] = 0$ whereas $S[\mathbb{I}/d] = \log d$ for a maximally mixed state. VN entropy is a monotone function of the purity, and viceversa.

Exercise 4 Evaluate purity and VN entropy of the partial traces of the state $|\psi\rangle\rangle = \cos \phi |01\rangle\rangle + \sin \phi |10\rangle\rangle$.

Exercise 5 Prove that for any pure bipartite state the entropies of the partial traces are equal, though the two density operators need not to be equal.

Exercise 6 Take a single-qubit state with density operator expressed in terms of the Pauli matrices $\varrho = \frac{1}{2}(\mathbb{I} + r_1\sigma_1 + r_2\sigma_2 + r_3\sigma_3)$ (Bloch sphere representation), $r_k = \text{Tr}[\varrho\sigma_k]$, and prove that the Bloch vector (r_1, r_2, r_3) should satisfies $r_1^2 + r_2^2 + r_3^2 \leq 1$ for ϱ to be a density operator.

3 Quantum measurements

In this section we put the postulates of standard quantum measurement theory under closer scrutiny. We start with some formal considerations and end up with a reformulation suitable for the description of any measurement performed on a quantum system, including those involving external systems or a noisy environment [2,3].

Let us start by reviewing the postulate of standard quantum measurement theory in a pedantic way, i.e. by expanding Postulate 2; ϱ denotes the state of the system before the measurement.

- [2.1] Any observable quantity is associated to a Hermitian operator X with spectral decomposition $X = \sum_x x |x\rangle\langle x|$. The eigenvalues are real and we assume for simplicity that they are nondegenerate. A measurement of X yields one of the eigenvalues x as possible outcomes.
- [2.2] The eigenvectors of X form a basis for the Hilbert space. The projectors $P_x = |x\rangle\langle x|$ span the entire Hilbert space, $\sum_x P_x = \mathbb{I}$.
- [2.3] The projectors P_x are orthogonal $P_x P_{x'} = \delta_{xx'} P_x$. It follows that $P_x^2 = P_x$ and thus that the eigenvalues of any projector are 0 and 1.

- [2.4] (Born rule) The probability that a particular outcome is found as the measurement result is

$$p_x = \text{Tr} [P_x \varrho P_x] = \text{Tr} [\varrho P_x^2] \stackrel{\star}{=} \text{Tr} [\varrho P_x] .$$

- [2.5] (Reduction rule) The state after the measurement (reduction rule or projection postulate) is

$$\varrho_x = \frac{1}{p_x} P_x \varrho P_x ,$$

if the outcome is x .

- [2.6] If we perform a measurement but we do not record the results, the post-measurement state is given by $\tilde{\varrho} = \sum_x p_x \varrho_x = \sum_x P_x \varrho P_x$.

The formulations [2.4] and [2.5] follow from the formulations for pure states, upon invoking the existence of a purification:

$$\begin{aligned} p_x &= \text{Tr}_{AB} [P_x \otimes \mathbb{I}_B |\psi_{AB}\rangle\rangle\langle\langle\psi_{AB}| P_x \otimes \mathbb{I}_B] = \text{Tr}_{AB} [|\psi_{AB}\rangle\rangle\langle\langle\psi_{AB}| P_x^2 \otimes \mathbb{I}_B] \\ &= \text{Tr}_A [\varrho_A P_x^2] \end{aligned} \quad (4)$$

$$\begin{aligned} \varrho_{Ax} &= \frac{1}{p_x} \text{Tr}_B [P_x \otimes \mathbb{I}_B |\psi_{AB}\rangle\rangle\langle\langle\psi_{AB}| P_x \otimes \mathbb{I}_B] = \frac{1}{p_x} P_x \text{Tr}_B [|\psi_{AB}\rangle\rangle\langle\langle\psi_{AB}|] P_x \\ &= \frac{1}{p_x} P_x \varrho_A P_x . \end{aligned} \quad (5)$$

The message conveyed by these postulates is that we can only predict the spectrum of the possible outcomes and the probability that a given outcome is obtained. On the other hand, the measurement process is random, and we cannot predict the actual outcome of each run. Independentl on its purity, a density operator ϱ does not describe the state of a single system, but rather an ensemble of identically prepared systems. If we perform the same measurement on each member of the ensemble we can predict the possible results and the probability with which they occur but we cannot predict the result of individual measurement (except when the probability of a certain outcome is either 0 or 1).

3.1 Probability operator-valued measure and detection operators

The set of postulates [2.*] may be seen as a set of recipes to generate probabilities and post-measurement states. We also notice that the number of possible outcomes is limited by the number of terms in the orthogonal resolution of identity, which itself cannot be larger than the dimensionality of the Hilbert space. It would however be often desirable to have more outcomes than the dimension of the Hilbert space while keeping positivity and normalization of probability distributions. In this section will show that this is formally possible, upon relaxing the assumptions on the mathematical objects describing the measurement, and replacing them with more flexible ones, still obtaining a meaningful prescription to generate probabilities. Then, in the next sections we will show that there are physical processes that fit with this generalized description, and that actually no revision of the postulates is needed, provided that the degrees of freedom of the measurement apparatus are taken into account.

The Born rule is a prescription to generate probabilities: its textbook form is the right term of the starred equality in [2.4]. However, the form on the left term has the merit to underline that in order to generate a probability it sufficient if the P_x^2 is a positive operator. In fact, we do not need to require that the set of the P_x 's are projectors, nor we need the positivity of the underlying P_x operators. So, let us consider

the following generalization: we introduce a set of positive operators $\Pi_x \geq 0$, which are the generalization of the P_x and use the prescription $p_x = \text{Tr}[\varrho \Pi_x]$ to generate probabilities. Of course, we want to ensure that this is a true probability distribution, i.e. normalized, and therefore require that $\sum_x \Pi_x = \mathbb{I}$, that is the positive operators still represent a resolution of the identity, as the set of projectors over the eigenstates of a selfadjoint operator. We will call a decomposition of the identity in terms of positive operators $\sum_x \Pi_x = \mathbb{I}$ a *probability operator-valued measure* (POVM) and $\Pi_x \geq 0$ the elements of the POVM.

Let us denote the operators giving the post-measurement states (as in [2.5]) by M_x . We refer to them as to the *detection operators*. As noted above, they are no longer constrained to be projectors. Actually, they may be any operator with the constraint, imposed by [2.4] i.e. $p_x = \text{Tr}[M_x \varrho M_x^\dagger] = \text{Tr}[\varrho \Pi_x]$. This tells us that the POVM elements have the form $\Pi_x = M_x^\dagger M_x$ which, by construction, individuate a set of a positive operators. There is a residual freedom in designing the post-measurement state. In fact, since Π_x is a positive operator $M_x = \sqrt{\Pi_x}$ exists and satisfies the constraint, as well as any operator of the form $M_x = U_x \sqrt{\Pi_x}$ with U_x unitary. This is the most general form of the detection operators satisfying the constraint $\Pi_x = M_x^\dagger M_x$ and corresponds to their polar decomposition. The POVM elements determine the absolute values leaving the freedom of choosing the unitary part.

Overall, the detection operators M_x represent a generalization of the projectors P_x , while the POVM elements Π_x generalize P_x^2 . The postulates for quantum measurements may be reformulated as follows

- [II.1] Observable quantities are associated to POVMs, i.e. decompositions of identity $\sum_x \Pi_x = \mathbb{I}$ in terms of positive $\Pi_x \geq 0$ operators. The possible outcomes are the eigenvalues x of the POVM elements.
- [II.2] The elements of a POVM are positive operators expressible as $\Pi_x = M_x^\dagger M_x$ where the detection operators M_x are generic operators with the only constraint $\sum_x M_x^\dagger M_x = \mathbb{I}$.
- [II.3] (Born rule) The probability that a particular outcome is found as the measurement result is $p_x = \text{Tr}[M_x \varrho M_x^\dagger] = \text{Tr}[\varrho M_x^\dagger M_x] = \text{Tr}[\varrho \Pi_x]$.
- [II.4] (Reduction rule) The state after the measurement is $\varrho_x = \frac{1}{p_x} M_x \varrho M_x^\dagger$ if the outcome is x .
- [II.5] If we perform a measurement but we do not record the results, the post-measurement state is given by $\tilde{\varrho} = \sum_x p_x \varrho_x = \sum_x M_x \varrho M_x^\dagger$.

Since orthogonality is no longer a requirement, the number of elements of a POVM has no restrictions and so the number of possible outcomes from the measurement. The above formulation generalizes both the Born rule and the reduction rule, and says that any set of detection operators satisfying [II.2] corresponds to a legitimate operations leading to a proper probability distribution and to a set of post-measurement states. This scheme is referred to as a *generalized measurement*.

Of course, up to this point, this is just a formal mathematical generalization of the standard description of measurements given in textbook quantum mechanics, and few questions naturally arise: Do generalized measurements describe physically realizable measurements? How they can be implemented? And if this is the case, does it means that standard formulation is too restrictive or wrong? To all these questions an answer will be provided by the following sections where we state and prove the Naimark Theorem, and discuss few examples of measurements described by POVMs.

3.2 The Naimark theorem

The Naimark theorem basically says that any generalized measurement satisfying [II.*] may be viewed as a standard measurement defined by [2.*] in a larger Hilbert

space, and conversely, any standard measurement involving more than one physical system may be described as a generalized measurement on one of the subsystems. In other words, if we focus attention on a portion of a composite system where a standard measurement takes place, then the statistics of the outcomes and the post-measurement states of the subsystem may be obtained with the tools of generalized measurements. Overall, we have

Theorem 4 (Naimark) *For any given POVM $\sum_x \Pi_x = \mathbb{I}$, $\Pi_x \geq 0$ on a Hilbert space H_A there exists a Hilbert space H_B , a state $\varrho_B = |\omega_B\rangle\langle\omega_B| \in L(H_B)$, a unitary operation $U \in L(H_A \otimes H_B)$, $UU^\dagger = U^\dagger U = \mathbb{I}$, and a projective measurement P_x , $P_x P'_x = \delta_{xx'} P_x$ on H_B such that $\Pi_x = \text{Tr}_B[\mathbb{I} \otimes \varrho_B U^\dagger \mathbb{I} \otimes P_x U]$. The setup is referred to as a Naimark extension of the POVM. Conversely, any measurement scheme where the system is coupled to an ancilla and, after evolution, a projective measurement is performed on the ancilla may be seen as the Naimark extension of a POVM, i.e. one may write the Born rule $p_x = \text{Tr}[\varrho_A \Pi_x]$ and the reduction rule $\varrho_A \rightarrow \varrho_{Ax} = \frac{1}{p_x} M_x \varrho_A M_x^\dagger$ at the level of the system only, in terms of the POVM elements $\Pi_x = \text{Tr}_B[\mathbb{I} \otimes \varrho_B U^\dagger \mathbb{I} \otimes P_x U]$ and the detection operators $M_x|\varphi_A\rangle = \langle x|U|\varphi_A, \omega_B\rangle$.*

Let us start with the second part of the theorem, and look at what happens when we couple the system under investigation to an additional system, usually referred to as ancilla (or apparatus), let them evolve, and then perform a projective measurement on the ancilla. This kind of setup is schematically depicted in Figure 1. The Hilbert space of the overall system is $H_A \otimes H_B$, and we assume that the system and the ancilla are initially independent on each other, i.e. the global initial preparation is $R = \varrho_A \otimes \varrho_B$. We also assume that the ancilla is prepared in the pure state $\varrho_B = |\omega_B\rangle\langle\omega_B|$ since this is always possible, upon a suitable purification of the ancilla degrees of freedom, i.e. by suitably enlarging the ancilla Hilbert space. Our aim is to obtain information about the system by measuring an observable X on the ancilla. This is done after the system-ancilla interaction described by the unitary operation U . According to the Born rule the probability of the outcomes is given by

$$p_x = \text{Tr}_{AB} [U \varrho_A \otimes \varrho_B U^\dagger \mathbb{I} \otimes |x\rangle\langle x|] = \text{Tr}_A \left[\varrho_A \underbrace{\text{Tr}_B [\mathbb{I} \otimes \varrho_B U^\dagger \mathbb{I} \otimes |x\rangle\langle x| U]}_{\Pi_x} \right]$$

where the set of operators $\Pi_x = \text{Tr}_B [\mathbb{I} \otimes \varrho_B U^\dagger \mathbb{I} \otimes |x\rangle\langle x| U] = \langle\omega_B|U^\dagger \mathbb{I} \otimes P_x U|\omega_B\rangle$ is the object that would permit to write the Born rule at the level of the subsystem A , i.e. it is our candidate POVM.

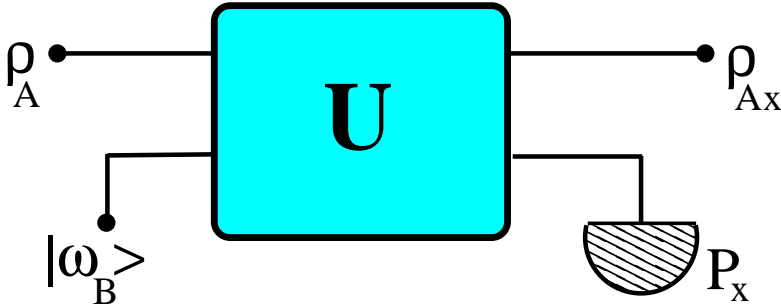


Fig. 1. Schematic diagram of a generalized measurement. The system of interest is coupled to an ancilla prepared in a known state $|\omega_B\rangle$ by the unitary evolution U , and then a projective measurement is performed on the ancilla.

In order to prove this, let us define the operators $M_x \in L(H_A)$ by their action on the generic vector in H_A

$$M_x|\varphi_A\rangle = \langle x|U|\varphi_A, \omega_B\rangle$$

where $|\varphi_A, \omega_B\rangle = |\varphi_A\rangle \otimes |\omega_B\rangle$ and the $|x\rangle$'s are the orthogonal eigenvectors of X . Using the decomposition of $\varrho_A = \sum_k \lambda_k |\psi_k\rangle\langle\psi_k|$ onto its eigenvectors the probability of the outcomes can be rewritten as

$$\begin{aligned} p_x &= \text{Tr}_{AB} [U \varrho_A \otimes \varrho_B U^\dagger \mathbb{I} \otimes |x\rangle\langle x|] = \sum_k \lambda_k \text{Tr}_{AB} [U |\psi_k, \omega_B\rangle\langle\omega_B, \psi_k| U^\dagger \mathbb{I} \otimes |x\rangle\langle x|] \\ &= \sum_k \lambda_k \text{Tr}_A [\langle x|U |\psi_k, \omega_B\rangle\langle\omega_B, \psi_k| U^\dagger |x\rangle] = \sum_k \lambda_k \text{Tr}_A [M_x |\psi_k\rangle\langle\psi_k| M_x^\dagger] \\ &= \text{Tr}_A [M_x \varrho_A M_x^\dagger] = \text{Tr}_A [\varrho_A M_x^\dagger M_x], \end{aligned} \quad (6)$$

which shows that $\Pi_x = M_x^\dagger M_x$ is indeed a positive operator $\forall x$. Besides, for any vector $|\varphi_A\rangle$ in H_A we have

$$\begin{aligned} \langle\varphi_A| \sum_x M_x^\dagger M_x |\varphi_A\rangle &= \sum_x \langle\langle\omega_B, \varphi_A| U^\dagger |x\rangle\langle x| U |\varphi_A, \omega_B\rangle\rangle \\ &= \langle\langle\omega_B, \varphi_A| U^\dagger U |\varphi_A, \omega_B\rangle\rangle = 1, \end{aligned} \quad (7)$$

and since this is true for any $|\varphi_A\rangle$ we have $\sum_x M_x^\dagger M_x = \mathbb{I}$. Putting together Eqs. (6) and (7) we have that the set of operators $\Pi_x = M_x^\dagger M_x$ is a POVM, with detection operators M_x . In turn, the conditional state of the system A , after having observed the outcome x , is given by

$$\begin{aligned} \varrho_{Ax} &= \frac{1}{p_x} \text{Tr}_B [U \varrho_A \otimes |\omega_B\rangle\langle\omega_B| U^\dagger \mathbb{I} \otimes P_x] = \frac{1}{p_x} \sum_k \lambda_k \langle x|U |\psi_k, \omega_B\rangle\langle\omega_B, \psi_k| U^\dagger |x\rangle \\ &= \frac{1}{p_x} M_x \varrho_A M_x^\dagger \end{aligned} \quad (8)$$

This is the half of the Naimark theorem: if we couple our system to an ancilla, let them evolve and perform the measurement of an observable on the ancilla, which projects the ancilla on a basis in H_B , then this procedure also modify the system. The transformation needs not to be a projection. Rather, it is adequately described by a set of detection operators which realizes a POVM on the system Hilbert space. Overall, the meaning of the above proof is twofold: on the one hand we have shown that there exists realistic measurement schemes which are described by POVMs when we look at the system only. At the same time, we have shown that the partial trace of a spectral measure is a POVM, which itself depends on the projective measurement performed on the ancilla, and on its initial preparation. Finally, we notice that the scheme of Figure 1 provides a general model for any kind of detector with internal degrees of freedom.

Let us now address the converse problem: given a set of detection operators M_x which realizes a POVM $\sum_x M_x^\dagger M_x = \mathbb{I}$, is this the system-only description of an indirect measurement performed a larger Hilbert space? In other words, there exists a Hilbert space H_B , a state $\varrho_B = |\omega_B\rangle\langle\omega_B| \in L(H_B)$, a unitary $U \in L(H_A \otimes H_B)$, and a projective measurement $P_x = |x\rangle\langle x|$ in H_B such that $M_x|\varphi_A\rangle = \langle x|U|\varphi_A, \omega_B\rangle$ holds for any $|\varphi_A\rangle \in H_A$ and $\Pi_x = \langle\omega_B|U^\dagger \mathbb{I} \otimes P_x U|\omega_B\rangle$? The answer is positive and we will provide a constructive proof. Let us take H_B with dimension equal to the number of detection operators and of POVM elements, and choose a basis $|x\rangle$ for H_B ,

which in turn individuates a projective measurement. Then we choose an arbitrary state $|\omega_B\rangle \in H_B$ and define the action of an operator U as

$$U|\varphi_A\rangle \otimes |\omega_B\rangle = \sum_x M_x |\varphi_A\rangle \otimes |x\rangle$$

where $|\varphi_A\rangle \in H_A$ is arbitrary. The operator U preserves the scalar product

$$\langle\langle\omega_B, \varphi'_A|U^\dagger U|\varphi_A, \omega_B\rangle\rangle = \sum_{xx'} \langle\varphi'_A|M_{x'}^\dagger M_x|\varphi_A\rangle \langle x'|x\rangle = \sum_x \langle\varphi'_A|M_{x'}^\dagger M_x|\varphi_A\rangle = \langle\varphi'_A|\varphi_A\rangle$$

and so it is unitary in the one-dimensional subspace spanned by $|\omega_B\rangle$. Besides, it may be extended to a full unitary operator in the global Hilbert space $H_A \otimes H_B$, eg it can be the identity operator in the subspace orthogonal to $|\omega_B\rangle$. Finally, for any $|\varphi_A\rangle \in H_A$, we have

$$\langle x|U|\varphi_A, \omega_B\rangle\rangle = \sum_{x'} M_{x'} |\varphi_A\rangle \langle x|x'\rangle = M_x |\varphi_A\rangle,$$

and

$$\langle\varphi_A|\Pi_x|\varphi_A\rangle = \langle\varphi_A|M_x^\dagger M_x|\varphi_A\rangle = \langle\langle\omega_B, \varphi_A|U^\dagger \mathbb{I} \otimes P_x U|\varphi_A, \omega_B\rangle\rangle,$$

that is, $\Pi_x = \langle\omega_B|U^\dagger \mathbb{I} \otimes P_x U|\omega_B\rangle$. \square

This completes the proof of the Naimark theorem, which asserts that there is a one-to-one correspondence between POVM and indirect measurements of the type describe above. In other words, an indirect measurement may be seen as the physical implementation of a POVM and any POVM may be realized by an indirect measurement.

The emerging picture is thus the following: In measuring a quantity of interest on a physical system one generally deals with a larger system that involves additional degrees of freedom, besides those of the system itself. These additional physical entities are globally referred to as the apparatus or the ancilla. As a matter of fact, the measured quantity may be always described by a standard observable, however on a larger Hilbert space describing both the system and the apparatus. When we trace out the degrees of freedom of the apparatus we are generally left with a POVM rather than a PVM. Conversely, any conceivable POVM, i.e. a set of positive operators providing a resolution of identity, describe a generalized measurement, which may be always implemented as a standard measurement in a larger Hilbert space.

Before ending this Section, few remarks are in order:

- R1** The possible Naimark extensions are actually infinite, corresponding to the intuitive idea that there are infinite ways, with an arbitrary number of ancillary systems, of measuring a given quantity. The construction reported above is sometimes referred to as the *canonical extension* of a POVM. The Naimark theorem just says that an implementation in terms of an ancilla-based indirect measurement is always possible, but of course the actual implementation may be different from the canonical one.
- R2** The projection postulate described at the beginning of this section, the scheme of indirect measurement, and the canonical extension of a POVM have in common the assumption that a nondemolitive detection scheme takes place, in which the system after the measurement has been modified, but still exists. This is sometimes referred to as a *measurement of the first kind* in textbook quantum mechanics. Conversely, in a demolitive measurement or *measurement of the second kind*, the system is destroyed during the measurement and it makes no sense of speaking of

the state of the system after the measurement. Notice, however, that for demolitive measurements on a field the formalism of generalized measurements provides the framework for the correct description of the state evolution. As for example, let us consider the detection of photons on a single-mode of the radiation field. A demolitive photodetector (as those based on the absorption of light) realizes, in ideal condition, the measurement of the number operator $a^\dagger a$ without leaving any photon in the mode. If $\varrho = \sum_{np} \varrho_{np} |n\rangle\langle p|$ is the state of the single-mode radiation field a photodetector of this kind gives a natural number n as output, with probability $p_n = \varrho_{nn}$, whereas the post-measurement state is the vacuum $|0\rangle\langle 0|$ independently on the outcome of the measurement. This kind of measurement is described by the orthogonal POVM $\Pi_n = |n\rangle\langle n|$, made by the eigenvectors of the number operator, and by the detection operator $M_n = |0\rangle\langle n|$. The proof is left as an exercise.

R3 We have formulated and proved the Naimark theorem in a restricted form, suitable for our purposes. It should be noticed that it holds in more general terms, as for example with extension of the Hilbert space given by direct sum rather than tensor product, and also relaxing the hypothesis [4].

3.2.1 Conditional states in generalized measurements

If we have a composite system and we perform a projective measurement on, say, subsystem A , the conditional state of the unmeasured subsystem B after the observation of the outcome x is given by Eq. (3), i.e. it is the partial trace of the projection of the state before the measurement on the eigenspace of the observed eigenvalue. One may wonder whether a similar results holds also when the measurement performed on the subsystem A is described by a POVM. The answer is positive and the proof may be given in two ways. The first is based on the observation that, thanks to the existence of a canonical Naimark extension, we may write the state of the global system after the measurement as

$$\varrho_{ABx} = \frac{1}{p_x} M_x \otimes \mathbb{I}_B \varrho_{AB} M_x^\dagger \otimes \mathbb{I}_B,$$

and thus the conditional state of subsystem B is the partial trace $\varrho_{Bx} = \text{Tr}_A[\varrho_{ABx}]$ i.e.

$$\varrho_{Bx} = \frac{1}{p_x} \text{Tr}_A[M_x \otimes \mathbb{I}_B \varrho_{AB} M_x^\dagger \otimes \mathbb{I}_B] = \frac{1}{p_x} \text{Tr}_A[\varrho_{AB} M_x^\dagger M_x \otimes \mathbb{I}_B] = \frac{1}{p_x} \text{Tr}_A[\varrho_{AB} \Pi_x \otimes \mathbb{I}_B],$$

where again we used the circularity of partial trace in the presence of factorized operators. A second proof may be offered invoking the Naimark theorem only to ensure the existence of an extension, i.e. a projective measurement on a larger Hilbert space $H_C \otimes H_A$, which reduces to the POVM after tracing over H_C . In formula, assuming that $P_x \in L(H_C \otimes H_A)$ is a projector and $\sigma \in L(H_C)$ a density operator

$$\begin{aligned} \varrho_{Bx} &= \frac{1}{p_x} \text{Tr}_{CA}[P_x \otimes \mathbb{I}_B \varrho_{AB} \otimes \sigma P_x \otimes \mathbb{I}_B] = \frac{1}{p_x} \text{Tr}_{CA}[\varrho_{AB} \otimes \sigma P_x \otimes \mathbb{I}_B] \\ &= \frac{1}{p_x} \text{Tr}_A[\varrho_{AB} \Pi_x \otimes \mathbb{I}_B]. \end{aligned}$$

3.3 Joint measurement of non commuting observables

A common statement about quantum measurements says that it is not possible to perform a joint measurement of two observables Q_A and P_A of a given system A

if they do not commute, i.e. $[Q_A, P_A] \neq 0$. This is related to the impossibility of finding any common set of projectors on the Hilbert space H_A of the system and to define a joint observable. On the other hand, a question arises on whether common projectors may be found in a larger Hilbert space, i.e. whether one may implement a joint measurement in the form of a generalized measurement. The answer is indeed positive [5,6]: This Section is devoted to describe the canonical implementation of joint measurements for pair of observables having a (nonzero) commutator $[Q_A, P_A] = c\mathbb{I} \neq 0$ proportional to the identity operator.

The basic idea is to look for a pair of commuting observables $[X_{AB}, Y_{AB}] = 0$ in a larger Hilbert space $H_A \otimes H_B$ which *trace* the observables P_A and Q_A , i.e. which have the same expectation values

$$\begin{aligned}\langle X_{AB} \rangle &\equiv \text{Tr}_{AB}[X_{AB} \varrho_A \otimes \varrho_B] = \text{Tr}_A[Q_A \varrho_A] \equiv \langle Q_A \rangle \\ \langle Y_{AB} \rangle &\equiv \text{Tr}_{AB}[Y_{AB} \varrho_A \otimes \varrho_B] = \text{Tr}_A[P_A \varrho_A] \equiv \langle P_A \rangle\end{aligned}\quad (9)$$

for any state $\varrho_A \in H_A$ of the system under investigation, and a fixed suitable preparation $\varrho_B \in H_B$ of the system B . A pair of such observables may be found upon choosing a replica system B , identical to A , and considering the operators

$$\begin{aligned}X_{AB} &= Q_A \otimes \mathbb{I}_B + \mathbb{I}_A \otimes Q_B \\ Y_{AB} &= P_A \otimes \mathbb{I}_B - \mathbb{I}_A \otimes P_B\end{aligned}\quad (10)$$

where Q_B and P_B are the analogue of Q_A and P_A for system B . The operators in Eq. (10), taken together a state $\varrho_B \in H_B$ satisfying

$$\text{Tr}_B[Q_B \varrho_B] = \text{Tr}_B[P_B \varrho_B] = 0, \quad (11)$$

fulfill the conditions in Eq. (9), i.e. realize a joint generalized measurement of the noncommuting observables Q_A and P_A . The operators X_{AB} and Y_{AB} are Hermitian by construction. Their commutator is given by

$$[X_{AB}, Y_{AB}] = [Q_A, P_A] \otimes \mathbb{I}_B - \mathbb{I}_A \otimes [Q_B, P_B] = 0. \quad (12)$$

Notice that the last equality, i.e. the fact that the two operators commute, is valid only if the commutator $[Q_A, P_A] = c\mathbb{I}$ is proportional to the identity. More general constructions are needed if this condition does not hold [7].

Since the $[X_{AB}, Y_{AB}] = 0$ the complex operator $Z_{AB} = X_{AB} + iY_{AB}$ is *normal* i.e. $[Z_{AB}, Z_{AB}^\dagger] = 0$. For normal operators the spectral theorem holds, and we may write

$$Z_{AB} = \sum_z z P_z \quad P_z = |z\rangle\langle z| \quad Z_{AB}|z\rangle = z|z\rangle \quad (13)$$

where $z \in \mathbb{C}$, and P_z are orthogonal projectors on the eigenstates $|z\rangle \equiv |z\rangle_{AB}$ of Z_{AB} . The set $\{P_z\}$ represents the common projectors individuating the joint observable Z_{AB} . Each run of the measurement returns a complex number, whose real and imaginary parts correspond to a sample of the X_{AB} and Y_{AB} values, aiming at sampling Q_A and P_A . The statistics of the measurement is given by

$$p_Z(z) = \text{Tr}_{AB}[\varrho_A \otimes \varrho_B P_z] = \text{Tr}_A[\varrho_A \Pi_z] \quad (14)$$

where the POVM Π_z is given by

$$\Pi_z = \text{Tr}_B[\mathbb{I}_A \otimes \varrho_B P_z]. \quad (15)$$

The mean values $\langle X_{AB} \rangle = \langle Q_A \rangle$ and $\langle Y_{AB} \rangle = \langle P_A \rangle$ are the correct ones by construction, where by saying "correct" we intend the mean values that one would have recorded by measuring the two observables Q_A and P_A separately in a standard (single) projective measurement on ϱ_A . On the other hand, the two marginal distributions

$$p_X(x) = \int dy p_Z(x + iy) \quad p_Y(y) = \int dx p_Z(x + iy),$$

need not to reproduce the distributions obtained in single measurements. In particular, for the measured variances $\langle \Delta X_{AB}^2 \rangle = \langle X_{AB}^2 \rangle - \langle X_{AB} \rangle^2$ and $\langle \Delta Y_{AB}^2 \rangle$ one obtains

$$\begin{aligned} \langle \Delta X_{AB}^2 \rangle &= \text{Tr} [(Q_A^2 \otimes \mathbb{I}_B + \mathbb{I}_A \otimes Q_B^2 + 2 Q_A \otimes Q_B) \varrho_A \otimes \varrho_B] - \langle Q_A \rangle^2 \\ &= \langle \Delta Q_A^2 \rangle + \langle Q_B \rangle^2 \\ \langle \Delta Y_{AB}^2 \rangle &= \langle \Delta P_A^2 \rangle + \langle P_B \rangle^2 \end{aligned} \quad (16)$$

where we have already taken into account that $\langle Q_B \rangle = \langle P_B \rangle = 0$. As it is apparent from Eqs. (16) the variances of X_{AB} and Y_{AB} are larger than those of the original, non commuting, observables Q_A and P_A .

Overall, we may summarize the emerging picture as follows: a joint measurement of a pair of non commuting observables corresponds to a generalized measurement and may be implemented as the measurement of a pair of commuting observables on an enlarged Hilbert space. Mean values are preserved whereas the non commuting nature of the original observables manifests itself in the broadening of the marginal distributions, i.e. as an additional noise term appears to both the variances. The uncertainty product may be written as

$$\begin{aligned} \langle \Delta X_{AB}^2 \rangle \langle \Delta Y_{AB}^2 \rangle &= \langle \Delta Q_A^2 \rangle \langle \Delta P_A^2 \rangle + \langle \Delta Q_A^2 \rangle \langle P_B^2 \rangle + \langle Q_B^2 \rangle \langle \Delta P_A^2 \rangle + \langle Q_B^2 \rangle \langle P_B^2 \rangle, \\ &\geq \frac{1}{4} |[Q_A, P_A]|^2 + \langle \Delta Q_A^2 \rangle \langle P_B^2 \rangle + \langle Q_B^2 \rangle \langle \Delta P_A^2 \rangle + \langle Q_B^2 \rangle \langle P_B^2 \rangle, \end{aligned} \quad (17)$$

where the last three terms are usually referred to as the *added noise* due to the joint measurement. If we perform a joint measurement on a minimum uncertainty state (MUS, see Appendix B) for a given pair of observables (e.g. a coherent state in the joint measurement of a pair of conjugated quadratures of the radiation field) and use a MUS also for the preparation of the replica system (e.g. the vacuum), then Eq. (17) rewrites as

$$\langle \Delta X_{AB}^2 \rangle \langle \Delta Y_{AB}^2 \rangle = |[Q_A, P_A]|^2. \quad (18)$$

This is four times the minimum attainable uncertainty product in the case of a measurement of a single observable (see Appendix B). In terms of rms' $\Delta X = \sqrt{\langle \Delta X^2 \rangle}$ we have a factor 2, which is usually referred to as the 3 dB of added noise in joint measurements. The experimental realization of joint measurements of non commuting observables has been carried out for conjugated quadratures of the radiation field in a wide range of frequencies ranging from radiowaves to the optical domain, see e.g. [8].

3.4 About the so-called Heisenberg principle

Let us start by quoting Wikipedia about the Heisenberg principle [9]

Published by Werner Heisenberg in 1927, the principle implies that it is impossible to simultaneously both measure the present position while "determining"

the future momentum of an electron or any other particle with an arbitrary degree of accuracy and certainty. This is not a statement about researchers' ability to measure one quantity while determining the other quantity. Rather, it is a statement about the laws of physics. That is, a system cannot be defined to simultaneously measure one value while determining the future value of these pairs of quantities. The principle states that a minimum exists for the product of the uncertainties in these properties that is equal to or greater than one half of the reduced Planck constant.

As is it apparent from the above formulation, the principle is about the precision achievable in the measurement of an observable and the disturbance introduced by the same measurement on the state under investigation, which, in turn, limits the precision of a subsequent measurement of the conjugated observable. The principle, which has been quite useful in the historical development of quantum mechanics, has been inferred from the analysis of the celebrated Heisenberg' gedanken experiments, and thus is heuristic in nature. However, since its mathematical formulation is related to that of the uncertainty relations (see Appendix B), it is often thought as a theorem following from the axiomatic structure of quantum mechanics. This is not the case: here we exploit the formalism of generalized measurements to provide an explicit example of a measurement scheme providing the maximum information about a given observable, i.e. the statistics of the corresponding PVM, while leaving the state under investigation in an eigenstate of the conjugated observable.

Let us consider the two noncommuting observables $[A, B] = c\mathbb{I}$ and the set of detection operators $M_a = |b\rangle\langle a|$ where $|a\rangle$ and $|b\rangle$ are eigenstates of A and B respectively, i.e. $A|a\rangle = a|a\rangle$, $B|b\rangle = b|b\rangle$. According to the Naimark theorem the set of operators $\{M_a\}$ describe a generalized measurement (e.g. an indirect measurement as the one depicted in Fig. 1) with statistics $p_a = \text{Tr}[\varrho \Pi_a]$ described by the POVM $\Pi_a = M_a^\dagger M_a = |a\rangle\langle a|$ and where the conditional states after the measurement are given by $\varrho_a = \frac{1}{p_a} M_a \varrho M_a^\dagger = |b\rangle\langle b|$. In other words, the generalized measurement described by the set $\{M_a\}$ has the same statistics of a Von-Neumann projective measurement of the observable A , and leave the system under investigating in an eigenstate of the observable B , thus *determining its future value with an arbitrary degree of accuracy and certainty* and contrasting the formulation of the so-called Heisenberg principle reported above. An explicit unitary realization of this kind of measurement for the case of position, as well as a detailed discussion on the exact meaning of the Heisenberg principle, and the tradeoff between precision and disturbance in a quantum measurement, may be found in [10].

3.5 The quantum roulette

Let us consider K projective measurements corresponding to K nondegenerate isospectral observables X_k , $k = 1, \dots, K$ in a Hilbert space H , and consider the following experiment. The system is sent to a detector which at random, with probability z_k , $\sum_k z_k = 1$, perform the measurement of the observable X_k . This is known as the quantum roulette since the observable to be measured is chosen at random, eg according to the outcome of a random generator like a roulette. The probability of getting the outcome x from the measurement of the observable X_k on a state $\varrho \in L(H)$ is given by $p_x^{(k)} = \text{Tr}[\varrho P_x^{(k)}]$, $P_x^{(k)} = |x\rangle_{kk}\langle x|$, and the overall probability of getting the outcome x from our experiment is given by

$$p_x = \sum_k z_k p_x^{(k)} = \sum_k z_k \text{Tr}[\varrho P_x^{(k)}] = \text{Tr}[\varrho \sum_k z_k P_x^{(k)}] = \text{Tr}[\varrho \Pi_x],$$

where the POVM describing the measurement is given by $\Pi_x = \sum_k z_k P_x^{(k)}$. This is indeed a POVM and not a projective measurement since

$$[\Pi_x, \Pi_{x'}] = \sum_{kk'} z_k z_{k'} [P_x^{(k)}, P_{x'}^{(k')}] \neq 0.$$

Again, we have a practical situation where POVMs naturally arise in order to describe the statistics of the measurement in terms of the Born rule and the system density operator. A Naimark extension for the quantum roulette may be obtained as follows. Let us consider an additional *probe* system described by the Hilbert space H_P of dimension K equal to the number of measured observables in the roulette, and the set of projectors $Q_x = \sum_k P_x^{(k)} \otimes |\theta_k\rangle\langle\theta_k|$ where $\{|\theta_k\rangle\}$ is a basis for H_P . Then, upon preparing the probe system in the superposition $|\omega_P\rangle = \sum_k \sqrt{z_k} |\theta_k\rangle$ we have that $p_x = \text{Tr}_{SP}[\varrho \otimes |\omega_P\rangle\langle\omega_P| Q_x]$ and, in turn, $\Pi_x = \text{Tr}_P[\mathbb{I}_S \otimes |\omega_P\rangle\langle\omega_P| Q_x] = \sum_k z_k P_x^{(k)}$. The state of the system after the measurement may be obtained as the partial trace

$$\begin{aligned} \varrho_x &= \frac{1}{p_x} \text{Tr}_P [Q_x \varrho \otimes |\omega_P\rangle\langle\omega_P| Q_x] \\ &= \frac{1}{p_x} \sum_k \sum_{k'} \text{Tr}_P \left[P_x^{(k)} \otimes |\theta_k\rangle\langle\theta_k| \varrho \otimes |\omega_P\rangle\langle\omega_P| P_x^{(k')} \otimes |\theta_{k'}\rangle\langle\theta_{k'}| \right] \\ &= \frac{1}{p_x} \sum_k z_k P_x^{(k)} \varrho P_x^{(k)}. \end{aligned}$$

Notice that the presented Naimark extension is not the canonical one.

Exercise 7 Prove that the operators Q_x introduced for the Naimark extension of the quantum roulette, are indeed projectors.

Exercise 8 Take a system made by a single qubit system and construct the canonical Naimark extension for the quantum roulette obtained by measuring the observables $\sigma_\alpha = \cos \alpha \sigma_1 + \sin \alpha \sigma_2$, where σ_1 and σ_2 are Pauli matrices and $\alpha \in [0, \pi]$ is chosen at random with probability $p(\alpha) = \pi^{-1}$.

4 Quantum operations

In this Section we address the dynamical evolution of quantum systems to see whether the standard formulation in terms of unitary evolutions needs a suitable generalization. This is indeed the case: we will introduce a generalized description and see how this reconciles with what we call Postulate 3 in the Introduction. We will proceed in close analogy with what we have done for states and measurements. We start by closely inspecting the physical motivations behind any mathematical description of quantum evolution, and look for physically motivated conditions that a map, intended to transform a quantum state into a quantum state, from now on a *quantum operation*, should satisfy to be admissible. This will lead us to the concept of complete positivity, which suitably generalizes the motivations behind unitarity. We then prove that any quantum operation may be seen as the partial trace of a unitary evolution in a larger Hilbert space, and illustrate a convenient form, the so-called Kraus or operator-sum representation, to express the action of a quantum operation on quantum states.

By quantum operation we mean a map $\varrho \rightarrow \mathcal{E}(\varrho)$ transforming a quantum state ϱ into another quantum state $\mathcal{E}(\varrho)$. The basic requirements on \mathcal{E} to describe a physically admissible operations are those captured by the request of unitarity in the standard formulation, i.e.

- Q1** The map is positive and trace-preserving, i.e. $\mathcal{E}(\varrho) \geq 0$ (hence selfadjoint) and $\text{Tr}[\mathcal{E}(\varrho)] = \text{Tr}[\varrho] = 1$. The last assumption may be relaxed to that of being trace non-increasing $0 \leq \text{Tr}[\mathcal{E}(\varrho)] \leq 1$ in order to include evolution induced by measurements (see below).
- Q2** The map is linear $\mathcal{E}(\sum_k p_k \varrho_k) = \sum_k p_k \mathcal{E}(\varrho_k)$, i.e. the state obtained by applying the map to the ensemble $\{p_k, \varrho_k\}$ is the ensemble $\{p_k, \mathcal{E}(\varrho_k)\}$.
- Q3** The map is completely positive (CP), i.e. besides being positive it is such that if we introduce an additional system, any map of the form $\mathcal{E} \otimes \mathbb{I}$ acting on the extended Hilbert space is also positive. In other words, we ask that the map is physically meaningful also when acting on a portion of a larger, composite, system. As we will see, this request is not trivial at all, i.e. there exist maps that are positive but not completely positive.

4.1 The operator-sum representation

This section is devoted to state and prove a theorem showing that a map is a quantum operation if and only if it is the partial trace of a unitary evolution in a larger Hilbert space, and provides a convenient form, the so-called Kraus decomposition or operator-sum representation [11,12], to express its action on quantum states.

Theorem 5 (Kraus) *A map \mathcal{E} is a quantum operation i.e. it satisfies the requirements **Q1-Q3** if and only if it is the partial trace of a unitary evolution on a larger Hilbert space with factorized initial condition or, equivalently, it possesses a Kraus decomposition i. e. its action may be represented as $\mathcal{E}(\varrho) = \sum_k M_k \varrho M_k^\dagger$ where $\{M_k\}$ is a set of operators satisfying $\sum_k M_k^\dagger M_k = \mathbb{I}$.*

Proof The first part of the theorem consists in assuming that $\mathcal{E}(\varrho)$ is the partial trace of a unitary operation in a larger Hilbert space and prove that it has a Kraus decomposition and, in turn, it satisfies the requirements **Q1-Q3**. Let us consider a physical system A prepared in the quantum state ϱ_A and another system B prepared in the state ϱ_B . A and B interact through the unitary operation U and we are interested in describing the effect of this interaction on the system A only, i.e. we are looking for the expression of the mapping $\varrho_A \rightarrow \varrho'_A = \mathcal{E}(\varrho_A)$ induced by the interaction. This may be obtained by performing the partial trace over the system B of the global AB system after the interaction, in formula

$$\begin{aligned} \mathcal{E}(\varrho_A) &= \text{Tr}_B [U \varrho_A \otimes \varrho_B U^\dagger] = \sum_s p_s \text{Tr}_B [U \varrho_A \otimes |\theta_s\rangle\langle\theta_s| U^\dagger] \\ &= \sum_{st} p_s \langle\varphi_t| U |\theta_s\rangle \varrho_A \langle\theta_s| U^\dagger |\varphi_t\rangle = \sum_k M_k \varrho_A M_k^\dagger \end{aligned} \quad (19)$$

where we have introduced the operator $M_k = \sqrt{p_s} \langle\varphi_t| U |\theta_s\rangle$, with the polyindex $k \equiv st$ obtained by a suitable ordering, and used the spectral decomposition of the density operator $\varrho_B = \sum_s p_s |\theta_s\rangle\langle\theta_s|$. Actually, we could have also assumed the additional system in a pure state $|\omega_B\rangle$, since this is always possible upon invoking a purification, i.e. by suitably enlarging the Hilbert space. In this case the elements in the Kraus decomposition of our map would have been written as $\langle\varphi_t| U |\omega_B\rangle$. The set of operators $\{M_k\}$ satisfies the relation

$$\sum_k M_k^\dagger M_k = \sum_{st} p_s \theta_s |U^\dagger |\varphi_t\rangle\langle\varphi_t| U |\theta_s\rangle = \sum_s p_s \langle\theta_s| U^\dagger U |\theta_s\rangle = \mathbb{I}.$$

Notice that the assumption of a factorized initial state is crucial to prove the existence of a Kraus decomposition and, in turn, the complete positivity. In fact, the dynamics

$\mathcal{E}(\varrho_A) = \text{Tr}_B [U \varrho_{AB} U^\dagger]$ resulting from assuming an initially correlated preparation ϱ_{AB} needs not be so. In this case, the dynamics can properly be defined only on a subset of initial states of the system. Of course, the map can be extended to all possible initial states by linearity, but the extension may not be physically realizable, i.e. may be not completely positive or even positive [13].

We now proceed to show that for map of the form (19) (Kraus decomposition) the properties **Q1-Q3** hold. Preservation of trace and of the Hermitian character, as well as linearity, are guaranteed by the very form of the map. Positivity is also ensured, since for any positive operator $O_A \in L(H_A)$ and any vector $|\varphi_A\rangle \in H_A$ we have

$$\begin{aligned} \langle \varphi_A | \mathcal{E}(O_A) | \varphi_A \rangle &= \langle \varphi_A | \sum_k M_k O_A M_k^\dagger | \varphi_A \rangle = \langle \varphi_A | \text{Tr}_B [U O_A \otimes \varrho_B U^\dagger] | \varphi_A \rangle \\ &= \text{Tr}_{AB} [U^\dagger | \varphi_A \rangle \langle \varphi_A | \otimes \mathbb{I} U O_A \otimes \varrho_B] \geq 0 \quad \forall O_A, \forall \varrho_B, \forall | \varphi_A \rangle. \end{aligned}$$

Therefore it remains to be proved that the map is completely positive. To this aim let us consider a positive operator $O_{AC} \in L(H_A \otimes H_C)$ and a generic state $|\psi_{AC}\rangle$ on the same enlarged space, and define

$$|\omega_k\rangle = \frac{1}{\sqrt{N_k}} M_k \otimes \mathbb{I}_C |\psi_{AC}\rangle \quad N_k = \langle \psi_{AC} | M_k^\dagger M_k \otimes \mathbb{I}_C | \psi_{AC} \rangle \geq 0.$$

Since O_{AC} is positive we have

$$\langle \psi_{AC} | (M_k^\dagger \otimes \mathbb{I}_C) O_{AC} (M_k \otimes \mathbb{I}_C) | \psi_{AC} \rangle = N_k \langle \omega_k | O_{AC} | \omega_k \rangle \geq 0$$

and therefore $\langle \psi_{AC} | \mathcal{E} \otimes \mathbb{I}_C (O_{AC}) | \psi_{AC} \rangle = \sum_k N_k \langle \omega_k | O_{AC} | \omega_k \rangle \geq 0$, which proves that for any positive O_{AC} also $\mathcal{E} \otimes \mathbb{I}_C (O_{AC})$ is positive for any choice of H_C , i.e. \mathcal{E} is a CP-map.

Let us now prove the second part of the theorem, i.e. we consider a map $\mathcal{E} : L(H_A) \rightarrow L(H_A)$ satisfying the requirements **Q1-Q3** and show that it may be written in the Kraus form and, in turn, that its action may be obtained as the partial trace of a unitary evolution in a larger Hilbert. We start by considering the state $|\varphi\rangle = \frac{1}{\sqrt{d}} \sum_k |\theta_k\rangle \otimes |\theta_k\rangle \in H_A \otimes H_A$ and define the operator $\varrho_{AA} = \mathcal{E} \otimes \mathbb{I}(|\varphi\rangle \langle \varphi|)$. From the complete positivity and trace preserving properties of \mathcal{E} we have that $\text{Tr}[\varrho_{AA}] = 1$, and $\varrho_{AA} \geq 0$, i.e. ϱ_{AA} is a density operator. Besides, this establishes a one-to-one correspondence between maps $L(H_A) \rightarrow L(H_A)$ and density operators in $L(H_A) \otimes L(H_A)$ which may be proved as follows: for any $|\psi\rangle = \sum_k \psi_k |\theta_k\rangle \in H_A$ define $|\tilde{\psi}\rangle = \sum_k \psi_k^* |\theta_k\rangle$ and notice that

$$\langle \tilde{\psi} | \varrho_{AA} | \tilde{\psi} \rangle = \frac{1}{d} \langle \tilde{\psi} | \sum_{kl} \mathcal{E}(|\theta_k\rangle \langle \theta_l|) \otimes |\theta_k\rangle \langle \theta_l| | \tilde{\psi} \rangle = \frac{1}{d} \sum_{kl} \psi_l^* \psi_k \mathcal{E}(|\theta_k\rangle \langle \theta_l|) = \frac{1}{d} \mathcal{E}(|\psi\rangle \langle \psi|),$$

where we used linearity to obtain the last equality. Then define the operators $M_k |\psi\rangle = \sqrt{d p_k} \langle \tilde{\psi} | \omega_k \rangle$, where $|\omega_k\rangle$ are the eigenvectors of $\varrho_{AA} = \sum_k p_k |\omega_k\rangle \langle \omega_k|$: this is a linear operator on H_A and we have

$$\sum_k M_k |\psi\rangle \langle \psi| M_k^\dagger = d \sum_k p_k \langle \tilde{\psi} | \omega_k \rangle \langle \omega_k | \tilde{\psi} \rangle = d \langle \tilde{\psi} | \varrho_{AA} | \tilde{\psi} \rangle = \mathcal{E}(|\psi\rangle \langle \psi|)$$

for all pure states. Using again linearity we have that $\mathcal{E}(\varrho) = \sum_k M_k \varrho M_k^\dagger$ also for any mixed state. It remains to be proved that a unitary extension exists, i.e. to prove that for any map on $L(H_A)$ which satisfies **Q1-Q3**, and thus possesses a Kraus decomposition, there exist: i) a Hilbert space H_B , ii) a state $|\omega_B\rangle \in H_B$, iii) a unitary

$U \in L(H_A \otimes H_B)$ such that $\mathcal{E}(\varrho_A) = \text{Tr}_B[U \varrho_A \otimes |\omega_B\rangle\langle\omega_B| U^\dagger]$ for any $\varrho_A \in L(H_A)$. To this aim we proceed as we did for the proof of the Naimark theorem, i.e. we take an arbitrary state $|\omega_B\rangle \in H_B$, and define an operator U through its action on the generic $\varphi_A \otimes |\omega_B\rangle \in H_A \otimes H_B$, $U |\varphi_A\rangle \otimes |\omega_B\rangle = \sum_k M_k |\varphi_A\rangle \otimes |\theta_k\rangle$, where the $|\theta_k\rangle$'s are a basis for H_B . The operator U preserves the scalar product

$$\langle\langle\omega_B, \varphi'_A|U^\dagger U|\varphi_A, \omega_B\rangle\rangle = \sum_{kk'} \langle\varphi'_A|M_{k'}^\dagger M_k|\varphi_A\rangle\langle\theta_{k'}|\theta_k\rangle = \sum_k \langle\varphi'_A|M_k^\dagger M_k|\varphi_A\rangle = \langle\varphi'_A|\varphi_A\rangle$$

and so it is unitary in the one-dimensional subspace spanned by $|\omega_B\rangle$. Besides, it may be extended to a full unitary operator in the global Hilbert space $H_A \otimes H_B$, eg it can be the identity operator in the subspace orthogonal to $|\omega_B\rangle$. Then, for any ϱ_A in H_A we have

$$\begin{aligned} \text{Tr}_B [U \varrho_A \otimes |\omega_B\rangle\langle\omega_B| U^\dagger] &= \sum_s p_s \text{Tr}_B [U |\psi_s\rangle\langle\psi_s| \otimes |\omega_B\rangle\langle\omega_B| U^\dagger] \\ &= \sum_{s k k'} p_s \text{Tr}_B [M_k |\psi_s\rangle\langle\psi_s| M_{k'}^\dagger \otimes |\theta_k\rangle\langle\theta_{k'}|] \\ &= \sum_{s k} p_s M_k |\psi_s\rangle\langle\psi_s| M_k^\dagger = \sum_k M_k \varrho_A M_k^\dagger \quad \square \end{aligned}$$

The Kraus decomposition of a quantum operation generalizes the unitary description of quantum evolution. Unitary maps are, of course, included and correspond to maps whose Kraus decomposition contains a single elements. The set of quantum operations constitutes a semigroup, i.e. the composition of two quantum operations is still a quantum operation:

$$\mathcal{E}_2(\mathcal{E}_1(\varrho)) = \sum_{k_1} M_{k_1}^{(1)} \mathcal{E}_2(\varrho) M_{k_1}^{(1)\dagger} = \sum_{k_1 k_2} M_{k_1}^{(1)} M_{k_2}^{(2)} \varrho M_{k_2}^{(2)\dagger} M_{k_1}^{(1)\dagger} = \sum_{\mathbf{k}} M_{\mathbf{k}} \varrho M_{\mathbf{k}}^\dagger,$$

where we have introduced the polyindex \mathbf{k} . Normalization is easily proved, since $\sum_{\mathbf{k}} M_{\mathbf{k}}^\dagger M_{\mathbf{k}} = \sum_{k_1 k_2} M_{k_2}^{(2)\dagger} M_{k_1}^{(1)\dagger} M_{k_1}^{(1)} M_{k_2}^{(2)} = \mathbb{I}$. On the other hand, the existence of inverse is not guaranteed: actually only unitary operations are invertible.

The Kraus theorem also allows us to have a unified picture of quantum evolution, either due to an interaction or to a measurement. In fact, the modification of the state in the both processes is described by a set of operators M_k satisfying $\sum_k M_k^\dagger M_k = \mathbb{I}$. In this framework, the Kraus operators of a measurement are what we have referred to as the detection operators of a POVM.

4.1.1 The dual map and the unitary equivalence

Upon writing the generic expectation value for the evolved state $\mathcal{E}(\varrho)$ and exploiting both linearity and circularity of trace we have

$$\langle X \rangle = \text{Tr}[\mathcal{E}(\varrho) X] = \sum_k \text{Tr}[M_k \varrho M_k^\dagger X] = \sum_k \text{Tr}[\varrho M_k^\dagger X M_k] = \text{Tr}[\varrho \mathcal{E}^\vee(X)],$$

where we have defined the dual map $\mathcal{E}^\vee(X) = \sum_k M_k^\dagger X M_k$ which represents the "Heisenberg picture" for quantum operations. Notice also that the elements of the Kraus decomposition $M_k = \langle\varphi_k|U|\omega_B\rangle$ depend on the choice of the basis used to

perform the partial trace. Change of basis cannot have a physical effect and this means that the set of operators

$$N_k = \langle \theta_k | U | \omega_B \rangle = \sum_s \langle \theta_k | \varphi_s \rangle \langle \varphi_s | U | \omega_B \rangle = \sum_s V_{ks} M_s,$$

where the unitary $V \in L(H_B)$ describes the change of basis, and the original set M_k actually describe the same quantum operations, i.e. $\sum_k N_k \varrho N_k^\dagger = \sum_k M_k \varrho M_k^\dagger$, $\forall \varrho$. The same can be easily proved for the system B prepared in mixed state. The origin of this degree of freedom stays in the fact that if the unitary U on $H_A \otimes H_B$ and the state $|\omega_B\rangle \in H_B$ realize an extension for the map $\mathcal{E} : L(H_A) \rightarrow L(H_A)$ then any unitary of the form $(\mathbb{I} \otimes V)U$ is a unitary extension too, with the same ancilla state. A quantum operation is thus identified by an equivalence class of Kraus decompositions. An interesting corollary is that any quantum operation on a given Hilbert space of dimension d may be generated by a Kraus decomposition containing at most d^2 elements, i.e. given a Kraus decomposition $\mathcal{E}(\varrho) = \sum_k M_k \varrho M_k^\dagger$ with an arbitrary number of elements, one may exploit the unitary equivalence and find another representation $\mathcal{E}(\varrho) = \sum_k N_k \varrho N_k^\dagger$ with at most d^2 elements.

4.2 The random unitary map and the depolarizing channel

A simple example of quantum operation is the random unitary map, defined by the Kraus decomposition $\mathcal{E}(\varrho) = \sum_k p_k U_k \varrho U_k^\dagger$, i.e. $M_k = \sqrt{p_k} U_k$ and $U_k U_k^\dagger = \mathbb{I}$. This map may be seen as the evolution resulting from the interaction of our system with another system of dimension equal to the number of elements in the Kraus decomposition of the map via the unitary V defined by $V|\psi_A\rangle \otimes |\omega_B\rangle = \sum_k \sqrt{p_k} U_k |\psi_A\rangle \otimes |\theta_k\rangle$, $|\theta_k\rangle$ being a basis for H_B which includes $|\omega_B\rangle$. If "we do not look" at the system B and trace out its degree of freedom the evolution of system A is governed by the random unitary map introduced above.

Exercise 9 Prove explicitly the unitarity of V .

The operator-sum representation of quantum evolutions have been introduced, and finds its natural application, for the description of propagation in noisy channels, i.e. the evolution resulting from the interaction of the system of interest with an external environment, which generally introduces noise in the system degrading its coherence. As for example, let us consider a qubit system (say, the polarization of a photon), on which we have encoded binary information according to a suitable coding procedure, traveling from a sender to a receiver. The propagation needs a physical support (say, an optical fiber) and this unavoidably leads to consider possible perturbations to our qubit, due to the interaction with the environment. The resulting open system dynamics is usually governed by a Master equation, i.e. the equation obtained by partially tracing the Schrodinger (Von Neumann) equation governing the dynamics of the global system, and the solution is expressed in form of a CP-map. For a qubit Q in a noisy environment a quite general description of the detrimental effects of the environment is the so-called depolarizing channel [1], which is described by the Kraus operator $M_0 = \sqrt{1-\gamma} \sigma_0$, $M_k = \sqrt{\gamma/3} \sigma_k$, $k = 1, 2, 3$, i.e.

$$\mathcal{E}(\varrho) = (1-\gamma)\varrho + \frac{\gamma}{3} \sum_k \sigma_k \varrho \sigma_k \quad 0 \leq \gamma \leq 1.$$

The depolarizing channel may be seen as the evolution of the qubit due to the interaction with a four-dimensional system through the unitary

$$V|\psi_Q\rangle \otimes |\omega_E\rangle = \sqrt{1-\gamma}|\psi_Q\rangle \otimes |\omega_E\rangle + \sqrt{\frac{\gamma}{3}} \sum_{k=1}^3 \sigma_k |\psi_Q\rangle \otimes |\theta_k\rangle,$$

$|\theta_k\rangle$ being a basis which includes $|\omega_E\rangle$. From the practical point view, the map describes a situation in which, independently on the underlying physical mechanism, we have a probability $\gamma/3$ that a perturbation described by a Pauli matrix is applied to the qubit. If we apply σ_1 we have the so-called spin-flip i.e. the exchange $|0\rangle \leftrightarrow |1\rangle$, whereas if we apply σ_3 we have the phase-flip, and for σ_2 we have a specific combination of the two effects. Since for any state of a qubit $\varrho + \sum_k \sigma_k \varrho \sigma_k = 2\mathbb{I}$ the action of the depolarizing channel may be written as

$$\mathcal{E}(\varrho) = (1-\gamma)\varrho + \frac{\gamma}{3}(2\mathbb{I} - \varrho) = \frac{2}{3}\gamma\mathbb{I} + (1 - \frac{4}{3}\gamma)\varrho = p\varrho + (1-p)\frac{\mathbb{I}}{2},$$

where $p = 1 - \frac{4}{3}\gamma$, i.e. $-\frac{1}{3} \leq p \leq 1$. In other words, we have that the original state ϱ is sent to a linear combination of itself and the maximally mixed state $\frac{\mathbb{I}}{2}$, also referred to as the depolarized state.

Exercise 10 Express the generic qubit state in Bloch representation and explicitly write the effect of the depolarizing channel on the Bloch vector.

Exercise 11 Show that the purity of a qubit cannot increase under the action of the depolarizing channel.

4.3 Transposition and partial transposition

The transpose $T(X) = X^T$ of an operator X is the conjugate of its adjoint $X^T = (X^\dagger)^* = (X^*)^\dagger$. Upon the choice of a basis we have $X = \sum_{nk} X_{nk} |\theta_n\rangle\langle\theta_k|$ and thus $X^T = \sum_{nk} X_{nk} |\theta_k\rangle\langle\theta_n| = \sum_{nk} X_{kn} |\theta_n\rangle\langle\theta_k|$. Transposition does not change the trace of an operator, neither its eigenvalues. Thus it transform density operators into density operators: $\text{Tr}[\varrho] = \text{Tr}[\varrho^T] = 1$ $\varrho^T \geq 0$ if $\varrho \geq 0$. As a positive, trace preserving, map it is a candidate to be a quantum operation. On the other hand, we will show by a counterexample that it fails to be completely positive and thus it does not correspond to physically admissible quantum operation. Let us consider a bipartite system formed by two qubits prepared in the state $|\varphi\rangle = \frac{1}{\sqrt{2}}|00\rangle + |11\rangle$. We denote by $\varrho^T = \mathbb{I} \otimes T(\varrho)$ the partial transpose of ϱ i.e. the operator obtained by the application of the transposition map to one of the two qubits. We have

$$\begin{aligned} (|\varphi\rangle\langle\varphi|)^T &= \frac{1}{2} \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}^T \\ &= \frac{1}{2} \left(|0\rangle\langle 0| \otimes |0\rangle\langle 0| + |1\rangle\langle 1| \otimes |1\rangle\langle 1| + |0\rangle\langle 1| \otimes |0\rangle\langle 1| + |1\rangle\langle 0| \otimes |1\rangle\langle 0| \right)^T \\ &= \frac{1}{2} \left(|0\rangle\langle 0| \otimes |0\rangle\langle 0| + |1\rangle\langle 1| \otimes |1\rangle\langle 1| + |0\rangle\langle 1| \otimes |1\rangle\langle 0| + |1\rangle\langle 0| \otimes |0\rangle\langle 1| \right) \\ &= \frac{1}{2} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{aligned}$$

Using the last expression it is straightforward to evaluate the eigenvalues of ϱ^T , which are $+\frac{1}{2}$ (multiplicity three) and $-\frac{1}{2}$. In other words $\mathbb{I} \otimes T$ is not a positive map and the transposition is not completely positive. Notice that for a factorized state of the form $\varrho_{AB} = \varrho_A \otimes \varrho_B$ we have $\mathbb{I} \otimes T(\varrho_{AB}) = \varrho_A \otimes \varrho_B^T \geq 0$ i.e. partial transposition preserves positivity in this case.

Exercise 12 Prove that transposition is not a CP-map by its action on any state of the form $|\varphi\rangle\rangle = \frac{1}{\sqrt{d}} \sum_k |\varphi_k\rangle \otimes |\theta_k\rangle$. Hint: the operator $\mathbb{I} \otimes T(|\varphi\rangle\rangle\langle\langle\varphi|) \equiv E$ is the so-called swap operator since it "exchanges" states as $E(|\psi\rangle_A \otimes |\varphi\rangle_B) = |\varphi\rangle_A \otimes |\psi\rangle_B$.

5 Conclusions

In this tutorial, we have addressed the postulates of quantum mechanics about states, measurements and operations. We have reviewed their modern formulation and introduced the basic mathematical tools: density operators, POVMs, detection operators and CP-maps. We have shown how they provide a suitable framework to describe quantum systems in interaction with their environment, and with any kind of measuring and processing devices. The connection with the standard formulation have been investigated in details building upon the concept of purification and the Theorems of Naimark and Stinespring/Kraus-Choi-Sudarshan.

The framework and the tools illustrated in this tutorial are suitable for the purposes of quantum information science and technology, a field which has fostered new experiments and novel views on the conceptual foundation of quantum mechanics, but has so far little impact on the way that it is taught. We hope to contribute in disseminating these notions to a larger audience, in the belief that they are useful for several other fields, from condensed matter physics to quantum biology.

I'm grateful to Konrad Banaszek, Maria Bondani, Mauro D'Ariano, Ivo P. Degiovanni, Marco Genoni, Marco Genovese, Paolo Giorda, Chiara Macchiavello, Sabrina Maniscalco, Alex Monras, Stefano Olivares, Jyrki Piilo, Alberto Porzio, Massimiliano Sacchi, Ole Steuernagel, and Bassano Vacchini for the interesting and fruitful discussions about foundations of quantum mechanics and quantum optics over the years. I would also like to thank Gerardo Adesso, Alessandra Andreoni, Alberto Barchielli, Rodolfo Bonifacio, Ilario Boscolo, Vlado Buzek, Berge Englert, Zdenek Hradil, Fabrizio Illuminati, Ludovico Lanz, Luigi Lugiato, Paolo Mataloni, Mauro Paternostro, Mladen Pavičić, Francesco Ragusa, Mario Rasetti, Mike Raymer, Jarda Řeháček, Salvatore Solimeno, and Paolo Tombesi.

References

1. M. Nielsen, E. Chuang, *Quantum Computation and Quantum Information*, (Cambridge University Press, 2000).
2. A. Peres, *Quantum Theory: concepts and methods*, (Kluwer Academic, Dordrecht, 1993).
3. J. Bergou, J. Mod. Opt. **57**, 160 (2010).
4. V. Paulsen, *Completely Bounded Maps and Operator Algebras* (Cambridge University Press, 2003).
5. E. Arthurs, J. L. Kelly, Bell. Syst. Tech. J. **44**, 725 (1965); J. P. Gordon, W. H. Louisell in *Physics of Quantum Electronics* (Mc-Graw-Hill, NY, 1966); E. Arthurs, M. S. Goodman, Phys. Rev. Lett. **60**, 2447 (1988).
6. H. P. Yuen, Phys. Lett. A **91**, 101 (1982).
7. E. Prugovečki, J. Phys. A **10**, 543 (1977).
8. N. G. Walker, J. E. Carol, Opt. Quantum Electr. **18**, 355 (1986); N. G. Walker, J. Mod. Opt. **34**, 16 (1987).

9. \protect\vrule width0pt\protect\href{http://en.wikipedia.org/wiki/Uncertainty_principle}{http://en.wikipedia.org/wiki/Uncertainty_principle}
10. M. Ozawa, Phys. Lett. A **299**, 17 (2002); Phys. Rev. A **67**, 042105 (2003); J. Opt. B **7**, S672 (2005).
11. J. Preskill, *Lectures notes for Physics 229: Quantum information and computation* available at www.theory.caltech.edu/~preskill/ph229/
12. Depending on the source and on the context, the theorem is known a Stinespring dilation theorem, or Kraus-Choi-Sudarshan theorem.
13. P. Pechukas, Phys. Rev. Lett. **73**, 1060 (1994).
14. R. Puri, *Mathematical methods of quantum optics* (Springer, Berlin, 2001).
15. K. E. Cahill, R. J. Glauber, Phys. Rev. **177**, 1857 (1969); **177**, 1882 (1969).

Further readings

1. I. Bengtsson, K. Życzkowski, *Geometry of Quantum States*, (Cambridge University Press, 2006).
2. Lectures and reports by C. M. Caves, available at <http://info.phys.unm.edu/~caves/>
3. P. Busch, M. Grabowski, P. J. Lahti, *Operational Quantum Mechanics*, Lect. Notes. Phys. **31**, (Springer, Berlin, 1995).
4. T. Heinosaari, M. Ziman, Acta Phys. Slovaca **58**, 487 (2008).
5. C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic Press, New York, 1976)
6. A.S. Holevo, *Statistical Structure of Quantum Theory*, Lect. Not. Phys **61**, (Springer, Berlin, 2001).
7. M. G. A. Paris, J. Rehacek (Eds.), *Quantum State Estimation* Lect. Notes Phys. **649**, (Springer, Berlin, 2004).
8. V. Gorini, A. Frigerio, M. Verri, A. Kossakowski, E. C. G. Sudarshan, Rep. Math. Phys. **13**, 149 (1978).
9. F. Buscemi, G. M. D'Ariano, and M. F. Sacchi, Phys. Rev. A **68**, 042113 (2003).
10. K. Banaszek, Phys. Rev. Lett. **86**, 1366 (2001).

A Trace and partial trace

The trace of an operator O is a scalar quantity equal to sum of diagonal elements in a given basis $\text{Tr}[O] = \sum_n \langle \varphi_n | O | \varphi_n \rangle$. The trace is invariant under any change of basis, as it is proved by the following chain of equalities

$$\begin{aligned} \sum_n \langle \theta_n | O | \theta_n \rangle &= \sum_{njk} \langle \theta_n | \varphi_k \rangle \langle \varphi_k | O | \varphi_j \rangle \langle \varphi_j | \theta_n \rangle = \sum_{njk} \langle \varphi_j | \theta_n \rangle \langle \theta_n | \varphi_k \rangle \langle \varphi_k | O | \varphi_j \rangle \\ &= \sum_{jk} \langle \varphi_j | \varphi_k \rangle \langle \varphi_k | O | \varphi_j \rangle = \sum_k \langle \varphi_k | O | \varphi_k \rangle, \end{aligned}$$

where we have suitably inserted and removed resolutions of the identity in terms of both basis $\{|\theta_n\rangle\}$ and $\{|\varphi_n\rangle\}$. As a consequence, using the basis of eigenvectors of O , $\text{Tr}[O] = \sum_n o_n$, o_n being the eigenvalues of O . Trace is a linear operation, i.e. $\text{Tr}[O_1 + O_2] = \text{Tr}[O_1] + \text{Tr}[O_2]$ and $\text{Tr}[\lambda O] = \lambda \text{Tr}[O]$ and thus $\partial \text{Tr}[O] = \text{Tr}[\partial O]$ for any derivation. The trace of any "ket-bra" $\text{Tr}[|\psi_1\rangle\langle\psi_2|]$ is obtained by "closing the sandwich" $\text{Tr}[|\psi_1\rangle\langle\psi_2|] = \langle\psi_2|\psi_1\rangle$; in fact upon expanding the two vectors in the same basis and taking the trace in that basis $\text{Tr}[|\psi_1\rangle\langle\psi_2|] = \sum_{nkl} \psi_{1k} \psi_{2l}^* \langle\theta_n|\theta_k\rangle \langle\theta_l|\theta_n\rangle = \sum_n \psi_{1n} \psi_{2n}^* = \langle\psi_2|\psi_1\rangle$. Other properties are summarized by the following theorem.

Theorem 6 *For the trace operation the following properties hold*

- i) *Given any pair of operators* $\text{Tr}[A_1 A_2] = \text{Tr}[A_2 A_1]$

ii) Given any set of operators A_1, \dots, A_N we $\text{Tr}[A_1 A_2 A_3 \dots A_N] = \text{Tr}[A_2 A_3 \dots A_N A_1] = \text{Tr}[A_3 A_4 \dots A_1 A_2] = \dots$ (circularity).

Proof : left as an exercise. \square

Notice that the "circularity" condition is essential to have property ii) i.e. $\text{Tr}[A_1 A_2 A_3] = \text{Tr}[A_2 A_3 A_1]$, but $\text{Tr}[A_1 A_2 A_3] \neq \text{Tr}[A_2 A_1 A_3]$

Partial traces $R_B \in L(H_B)$ $R_A \in L(H_A)$ of an operator R in $L(H_1 \otimes H_2)$ are defined accordingly as

$$R_B = \text{Tr}_A [R] = \sum_n {}_A \langle \varphi_n | R | \varphi_n \rangle_A \quad R_A = \text{Tr}_B [R] = \sum_n {}_B \langle \varphi_n | R | \varphi_n \rangle_B$$

and circularity holds only for single-system operators, e.g., if $R_1, R_2 \in L(H_A \otimes H_B)$, $A \in L(H_A)$, $B \in L(H_B)$

$$\begin{aligned} \text{Tr}_A [A \otimes \mathbb{I} R_1 R_2] &= \sum_n a_n \langle a_n | R_1 R_2 | a_n \rangle = \text{Tr}_A [R_1 R_2 A \otimes \mathbb{I}] \\ \text{Tr}_A [A \otimes B R_1 R_2] &= \sum_n a_n \langle a_n | \mathbb{I} \otimes B R_1 R_2 | a_n \rangle = \text{Tr}_A [\mathbb{I} \otimes B R_1 R_2 A \otimes \mathbb{I}] \\ &\neq \sum_n a_n \langle a_n | R_1 R_2 \mathbb{I} \otimes B | a_n \rangle = \text{Tr}_A [R_1 R_2 A \otimes B] \end{aligned}$$

Exercise 13 Consider a generic mixed state $\rho \in L(H \otimes H)$ and write the matrix elements of the two partial traces in terms of the matrix elements of ρ .

Exercise 14 Prove that also partial trace is invariant under change of basis.

B Uncertainty relations

Two non commuting observables $[X, Y] \neq 0$ do not admit common eigenvectors and thus it not possible to find common eigenprojectors and define a joint observable. Two non commuting observables are said to be incompatible or complementary, since they cannot assume definite values simultaneously. A striking consequence of this fact is that when we measure an observable X the precision of the measurement, as quantified by the variance $\langle \Delta X^2 \rangle = \langle X^2 \rangle - \langle X \rangle^2$, is influenced by the variance of any observable which is non commuting with X and cannot be made arbitrarily small. In order to determine the relationship between the variances of two noncommuting observables, one of which is measured on a given state $|\psi\rangle$, let us consider the two vectors

$$|\psi_1\rangle = (X - \langle X \rangle) |\psi\rangle \quad |\psi_2\rangle = (Y - \langle Y \rangle) |\psi\rangle,$$

and write explicitly the Schwartz inequality $\langle \psi_1 | \psi_1 \rangle \langle \psi_2 | \psi_2 \rangle \geq |\langle \psi_1 | \psi_2 \rangle|^2$, i.e. [14]

$$\langle \Delta X^2 \rangle \langle \Delta Y^2 \rangle \geq \frac{1}{4} \left[|\langle F \rangle|^2 + |\langle C \rangle|^2 \right] \geq \frac{1}{4} |\langle C \rangle|^2, \quad (20)$$

where $[X, Y] = iC$ and $F = XY - YX - 2\langle X \rangle \langle Y \rangle$. Ineq. (20) represents the uncertainty relation for the non commuting observables X and Y and it is usually presented in the form involving the second inequality. Uncertainty relations set a lower bound to the measured variance in the measurement of a single observable, say X , on a state with a fixed, intrinsic, variance of the complementary observable Y (see Section 3.3

for the relationship between the variance of two non commuting observables in a joint measurement). The uncertainty product is minimum when the two vectors $|\psi_1\rangle$ and $|\psi_2\rangle$ are parallel in the Hilbert space, i.e. $|\psi_1\rangle = -i\lambda|\psi_2\rangle$ where λ is a complex number. Minimum uncertainty states (MUS) for the pair of observables X, Y are thus the states satisfying

$$(X + i\lambda Y) |\psi\rangle = (\langle X \rangle + i\lambda \langle Y \rangle) |\psi\rangle.$$

If λ is real then $\langle F \rangle = 0$, i.e. the quantities X and Y are uncorrelated when the physical system is prepared in the state $|\psi\rangle$. If $|\lambda| = 1$ then $\langle \Delta X^2 \rangle = \langle \Delta Y^2 \rangle$ and the corresponding states are referred to as equal variance MUS. Coherent states of a single-mode radiation field [15] are equal variance MUS, e. g. for the pair of quadrature operators defined by $Q = \frac{1}{\sqrt{2}}(a^\dagger + a)$ and $P = \frac{i}{\sqrt{2}}(a^\dagger - a)$.